

BIROn - Birkbeck Institutional Research Online

Hahn, Ulrike and Hansen, J.U. and Olsson, E.J. (2020) Truth tracking performance of social networks: how connectivity and clustering can make groups less competent. *Synthese* 197 , pp. 1511-1541. ISSN 0039-7857.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/23842/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively



Truth tracking performance of social networks: how connectivity and clustering can make groups less competent

Ulrike Hahn¹ · Jens Ulrik Hansen² · Erik J. Olsson² 

Received: 6 December 2017 / Accepted: 5 September 2018
© The Author(s) 2018

Abstract

Our beliefs and opinions are shaped by others, making our social networks crucial in determining what we believe to be true. Sometimes this is for the good because our peers help us form a more accurate opinion. Sometimes it is for the worse because we are led astray. In this context, we address via agent-based computer simulations the extent to which patterns of connectivity within our social networks affect the likelihood that initially undecided agents in a network converge on a true opinion following group deliberation. The model incorporates a fine-grained and realistic representation of belief (opinion) and trust, and it allows agents to consult outside information sources. We study a wide range of network structures and provide a detailed statistical analysis concerning the exact contribution of various network metrics to collective competence. Our results highlight and explain the collective risks involved in an overly networked or partitioned society. Specifically, we find that 96% of the variation in collective competence across networks can be attributed to differences in amount of connectivity (average degree) and clustering, which are negatively correlated with collective competence. A study of bandwagon or “group think” effects indicates that both connectivity and clustering increase the probability that the network, wholly or partly, locks into a false opinion. Our work is interestingly related to Gerhard Schurz’s work on meta-induction and can be seen as broadly addressing a practical limitation of his approach.

Keywords Social networks · Truth · Network structure · Connectivity · Network metric · Collective competence · Veritistic value · Bandwagoning · Group think · Computer simulation · Gerhard Schurz

✉ Erik J. Olsson
Erik_J.Olsson@fil.lu.se

Extended author information available on the last page of the article

1 Introduction

Does your social network influence what you believe to be true? Most probably, the reader would answer this question in the affirmative. We may ask the further question: does the *structure* of your social network, i.e. the pattern of communication, influence what you believe to be true? This is perhaps less clear. The general question we address in this paper is how, if at all, network topology affects the group's ability to track truth.

There is an obvious commonsense or internet age answer to this question according to which the more connected a community of agents is, the better it will be at tracking truth. It would follow that the fully connected network, wherein everyone is connected to everyone else, maximizes the truth tracking ability of the group. Despite its intuitive appeal, the common sense answer has been undermined in a number of studies. This goes for agent-based models (e.g. Bala and Goyal 1998; Zollman 2007; Lazer and Freidman 2007) as well as for empirical studies (e.g. Mason et al. 2008; Jönsson et al. 2015). The bottom line is that having more connections in a group can be bad from a truth-tracking perspective, although several studies have found that it often speeds up a group's ability to converge on an opinion, be it true or false (e.g. Zollman 2007; Mason et al. 2008), and that the "less is more" effect depends on context (Frey and Šešelja 2018). What this suggests is that there are few simple truths in this area of research.

However, studies in the literature typically focus on only a few network topologies. For example, Mason et al. (2008) confine their attention to four network types. It would be desirable to look at a richer set of networks. Also, in terms of explaining the performance of various network structures, there is an emphasis on network density (the number of actual connections divided by the number of possible ones) at the expense of other network metrics (e.g. Zollman 2007), though as our study will show, density turns out to be a key factor.

Unfortunately, the by far most influential agent-based model for studying collective competence, introduced in Hegselmann and Krause (2002, 2006), lacks the flexibility necessary for studying the general effects of network structure. As Hegselmann has pointed out to us (personal communication), there are two ways to think about the H–K model in terms of network structure: (a) There is an underlying fully connected network, but only the links to agents whose opinions are sufficiently similar count for the updating procedure. (b) There is a fixed number of agents. Between these agents a process of dynamical networking is going on. In each period the actual network is given by linking all agents with opinions that are sufficiently similar, and updating is averaging over all linked opinions. Either way, it is not possible to study the effect of network structure independently of similarity of opinion. The H–K model, further, shares a limitation of the earlier DeGroot model in being "rigid in that agents do not adjust the weights they place on others' opinions over time" (Golub and Jackson 2010, p. 113).

Instead, we will use a Bayesian agent-based model called Laputa (e.g. Masterton and Olsson 2014; Olsson 2011; Olsson and Vallinder 2013; Vallinder and Olsson 2014). The rich and flexible Laputa framework (Douven and Kelp 2011) allows one to model agents that match the challenge of real-world information acquisition in two fundamental ways: (1) we receive evidence both from our own observation or

interaction with the world and through the testimony of others and (2) we do not come to either of these sources knowing accurately their reliability. Hence Laputa incorporates a mechanism for representing and updating both trust among agents in the network, in the sense of “perceived reliability”, i.e. the weights they place on others’ opinions over time, and for estimating the reliability of agent’s own inquiry.

While one would expect the topology of the social network constituted via the Laputa agents’ communication to affect their accuracy, it seems difficult to predict the exact impact given the potentially complex interactions between perceived reliability, communication, and evidence from the world. Actual simulations are thus required. Our study will focus on a comparatively large selection of networks and a number of well-known network metrics. Using regression analysis we identify the network characteristics that make unique contributions to collective competence (in Goldman’s 1999 sense), focusing on a scenario in which the agents are initially more or less undecided on the issue at hand.

Our work is interestingly related to Schurz’s seminal work on “meta-induction”—in the simplest case copying the method of the most successful network peer—and can be seen as broadly addressing a practical limitation of his approach. As Schurz points out, in a society characterized by division of cognitive labor indicators of trustworthiness of purported informants is of the utmost importance (Schurz, 2009, p. 201; see also Thorn and Schurz 2012). Schutz distinguishes between two sources of trustworthiness. One possibility is to understand trustworthiness as something entirely internal to society. Schurz calls this position “goal-internalism”. The other possibility is to conceive it, following Goldman (1999), as deriving from something external, namely objective reliability. This is the “goal-externalist” position recommended by Schurz. Now, as Schurz’s notes, meta-induction presupposes “an objective and consensual criterion of past successes in predictions (or actions based on predictions)” (ibid, p. 2018). In other words, the meta-inductive agent must know the track-records of the other agents. The problem is that this condition is arguably rarely satisfied in practice. What to do when it is not? This is where our study becomes relevant. As indicated in Collins et al. (2018), when a track-record is missing, people are happy to update their trust in a given source on the basis of message content. If the source says something expected, this tells in favor of the source’s reliability. If, by contrast, the source says something unexpected, it tells against the source’s reliability. While being firmly goal-externalist in the sense that outside sources are assumed to have an objective reliability, our model assumes that the network agents have access to this reliability only indirectly through the messages produced by the sources. Agents are thus forced to infer trustworthiness based on message content in the way described. Our model incorporated this idea in a full-fledged Bayesian framework. As we will see, the flip side of the coin is that group performance becomes negatively affected by connectivity and clustering to the extent that agents may be better off not communicating at all, relying only on their outside source.

2 Method

To appreciate the results reported here, it is necessary only that the reader grasp the broad features of the Laputa model. Consequently, this section conveys only the basic ideas behind the framework. More detailed expositions can be found in Olsson (2013) and Vallinder and Olsson (2014). Agents in the network are assumed to be concerned with answering the question whether p , where p is a proposition which can be true or false. Agents in the network start out with a certain degree of belief (credence) in p . What happens then is that the agents engage in inquiry and deliberation in the sense that they can receive information from an outside source or from network peers. This takes place in a number of rounds or steps representing an opportunity to receive information from inquiry or other agents, or transmit information to other agents.

The credence assigned to a proposition p by an agent α after inquiry and deliberation in a social network depends (among other things) on:

- reports from α 's outside source
- how many of α 's peers claim that p or not- p
- how often they do it
- α 's trust in (perceived reliability of) her peers

Thus an agent will be impressed by repeated information coming from many different sources, especially highly trusted ones.

Every model of a complex part of reality needs to be simplified and streamlined to be at all workable. Models of social network communication are no exceptions. As for Laputa, the following are assumed:

- At every round in deliberation, inquirers (outside sources) can communicate p , not- p or be silent
- Trust is modelled as a second order probability, i.e. as a credence in the reliability of the source
- Reports coming from different sources at the same time are viewed by receiving inquirers as independent
- Reports from outside sources are treated as independent

Olsson (2013) argues that the assumptions are justifiable from a dual process perspective. Thus, the constraint that agents in the model view each other as independent information sources can be viewed as a plausible default strategy attributed to system (process) 1 employed in the absence of concrete signs of trouble. The independence of outside sources is compatible with many plausible scenarios, such as the incoming information deriving from disjoint personal networks of the agents (Fig. 1).

Olsson (2013) also gives reason to think that the broad features of the model are in line with the influential Persuasive Argument Theory tradition in social psychology (for an overview, see Isenberg 1986). For instance, agents in Laputa polarize in the sense of Sunstein (2002).

As noted above, a feature of the Laputa model is that both degrees of belief (credences) and trust values are updated dynamically in the process of inquiry and deliberation. In both cases, updating takes place in adherence to the Bayesian principle of conditionalization on the new evidence. To illustrate, agent α 's new credence in p after hearing that p (or its negation, $\neg p$) from source σ is given by:

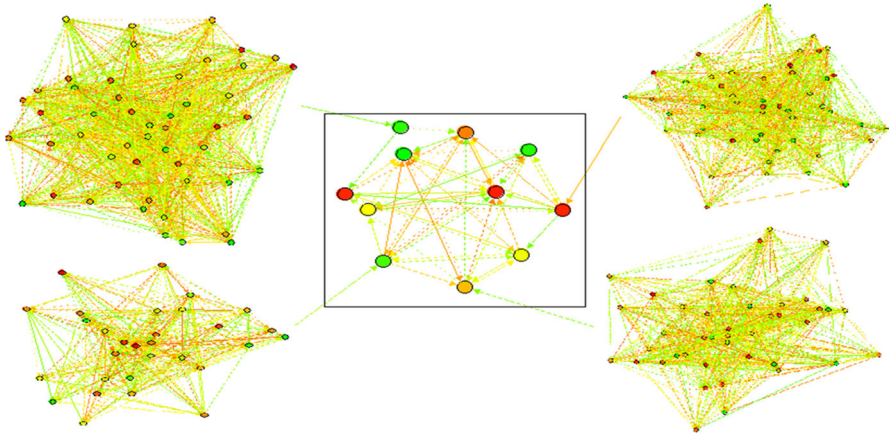


Fig. 1 Outside information deriving from disjoint personal networks

$$C_{\alpha}^t(p|S_{\sigma\alpha}^t(p)) = \frac{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t]}{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]}$$

$$C_{\alpha}^t(p|S_{\sigma\alpha}^t(\neg p)) = \frac{C_{\alpha}^t(p) \times E[\overline{\tau_{\sigma\alpha}^t}]}{C_{\alpha}^t(p) \times E[\overline{\tau_{\sigma\alpha}^t}] + C_{\alpha}^t(\neg p) \times E[\tau_{\sigma\alpha}^t]}$$

Here $C_{\alpha}^t(p)$ denotes the agent α 's credence in p at time t , and $E[\tau_{\sigma\alpha}^t]$ the expected value of the trust function assigned to source σ by agent α . As for the top equation, the new credence in p for agent α after receiving the information p from network peer σ is the old credence in p conditional on the fact that σ reported that p . This in turn can be reduced to the expression on the right-hand side which depends on the old credence for p ($\neg p$) as well as the (expected value) of the trust function which α associates with σ . A report that not- p ($\neg p$) is handled analogously (see bottom equation). For the function for updating trust and its derivation, see "Appendix C".

The underlying Bayesian machinery gives rise to some suggestive qualitative updating rules for credences and trust values (Table 1). A + sign means stronger credence (in the current direction), an up-arrow more trust, and so on. For example, a trusted source reporting an expected message leads to the receiver strengthening her current credence in the message as well as her trust in the source (upper left-most box). To take another example, a distrusted source delivering an unexpected message leads to the receiver strengthening her current credence but lowering her trust (lower right-most box), revealing that being distrusted in Laputa amounts to being viewed as a falsity-teller.

A network structure or topology is a particular kind of social arrangement. Our interest in this paper is in the truth-tracking properties of social arrangements as studied within social epistemology. In measuring the truth-tracking performance of

Table 1 Qualitative rules for updating credences and trust in Laputa

	Message expected	Neither nor	Message unexpected
Source trusted	+ (↑)	↑ ()	− (↓)
Neither nor	0 (↑)	0 ()	0 (↓)
Source distrusted	− (↑)	↓ ()	+ (↓)

a topology we follow Goldman (1999), specifically his theory of veritistic value (V-value, for short), in assuming that, ideally, an agent should have full belief in the truth. If it is true that it will rain, then an agent should fully believe that it will, i.e. assign credence 1 to that fact. If it is true that the Eurozone will collapse, then an agent should believe fully that it will, and so on (assuming, of course, that the agent cares about these propositions in the first place). More generally, inquirers are better off the closer they are to fulfilling this ideal, i.e. the closer their degree of belief in the truth is to full belief in the truth. So if it is true that it will rain, then an agent assigning credence 0.7 to that proposition is better off than an agent assigning only credence 0.6.

From this perspective, a network topology is epistemically advantageous to the extent that agents engaging in group deliberation constrained by that topology move closer to the truth on average. Thus, a network structure which is such that when agents allows it to govern their communication makes the agents more inclined to assign high credence to the truth is better than a network structure which does not have this property, or has it but to a lesser degree. In our simulations, we assume, by convention, that the proposition p is true and hence that its negation, not- p , is false. This also means that the collective accuracy of the agents in the simulation can be represented simply by the average degree of belief. There is a sizeable literature on how best to measure accuracy (Maher 1993; Joyce 1998; Fallis 2007; Kopec 2012), and, in particular, whether it requires the use of a so-called proper scoring rule, such as the squared error or “Brier score” (Brier 1950). This question can be set aside, because reporting the average degree of belief, and the increase of the average degree of belief, given the convention that the true value is always 1, will lead to the same answers vis-à-vis our central question as a monotonic transformation such as squaring the deviation of that mean to probability 1. Our interest lies with the effects of topology on collective competence and the network properties that mediate it. The presence or absence of such effects is unaffected by such transformations, and the regressions we conduct identify the same moderators using the absolute deviation and the squared error, varying only slightly in the absolute goodness-of-fit obtained. Consequently, we report absolute deviations between the average degree of belief in the network and the true value, or what has been referred to as veritistic value (Goldman 1999).¹

The Laputa model has been implemented in a computer program bearing the same name. Once a given network has been implemented, the Laputa program can run tens

¹ Differences between absolute and squared error only emerge when one considers not measures of collective accuracy, but individual accuracy, such as the mean individual error (see also Jönsson et al. 2015). This may affect the rank order of the topologies with respect to accuracy, but, once again, it does not affect the fact *that* topology influences accuracy. We pursue the differences between individual and collective competence in more detail elsewhere (Hahn et al. in preparation).

of thousands of simulations (group deliberations) using the same network structure. The program then outputs the average V-value and other useful statistical information.

Laputa is flexible in the sense that it allows for a number of parameters to be determined before running a simulation. In this study, we focus on a scenario in which all agents are initially more or less unsure about the truth of the proposition p . This is captured by having agents' initial credence in p selected from a normal distribution with expected value 0.5 and standard deviation 0.1. This means that when the Laputa simulator creates the initial state of a network, it picks the initial credences for the agents in the network from such a distribution. In other words, agents will, on average, start out with a credence of 0.5 in p , although some start out slightly lower and others slightly higher. This kind of scenario would be realistic for instance if the agents are deliberating on a new issue regarding which they have not yet reached a firm opinion. Note that there is no particular relationship between an agent's initial credence in p and the reliability of his or her outside source. The parameter values for the latter are described below.

As for the other parameters, we were careful to select distributions that can plausibly be said to capture a normal situation:

- (a) Agents engage in communication for some time but not indefinitely. Our simulations cover both medium and longer communicational activity (15 vs. 30 simulation rounds).
- (b) Agents rely on outside sources that are at least somewhat reliable, and they initially trust, to some extent, their sources and each other. Also, they don't have to be absolutely sure that they are right in order to communicate with their peers; it is sufficient that the credence is above a given threshold, called the communication threshold. In the simulations, parameter values for reliability of inquiry (= outside source), initial inquiry trust, initial peer trust and communication threshold were selected from a normal distribution with expected value 0.748 and standard deviation 0.098.
- (c) Finally, we assumed that agents reasonably often ask their outside sources and communicate their view given that their credence meets the requirement set by the communication threshold. Accordingly, the parameters inquiry chance and communication chance were selected from an interval distribution with expected value 0.5 and standard deviation 0.0289.

For example, the assumptions we have made would be reasonable in a case of jury deliberation in which the jurors, who do not know each other, are initially ignorant regarding the guilt of the defendant, assuming that the jurors are normal in terms of having a somewhat trusting, outspoken and reliable nature but varying in their level of activity regarding inquiry (which in this case can be viewed as involving consultation of their memory of the trial) as well as communication with other jurors. Since in a normal jury every juror can communicate with any other, this example would involve a fully connected network (see Fig. 2). The assumptions are also plausible for capturing online communication in an anonymous setting in which, as in the juror case, people do not know each other's true reliability, and in which participants are initially ignorant regarding the true answer to the underlying question. In this case, many network

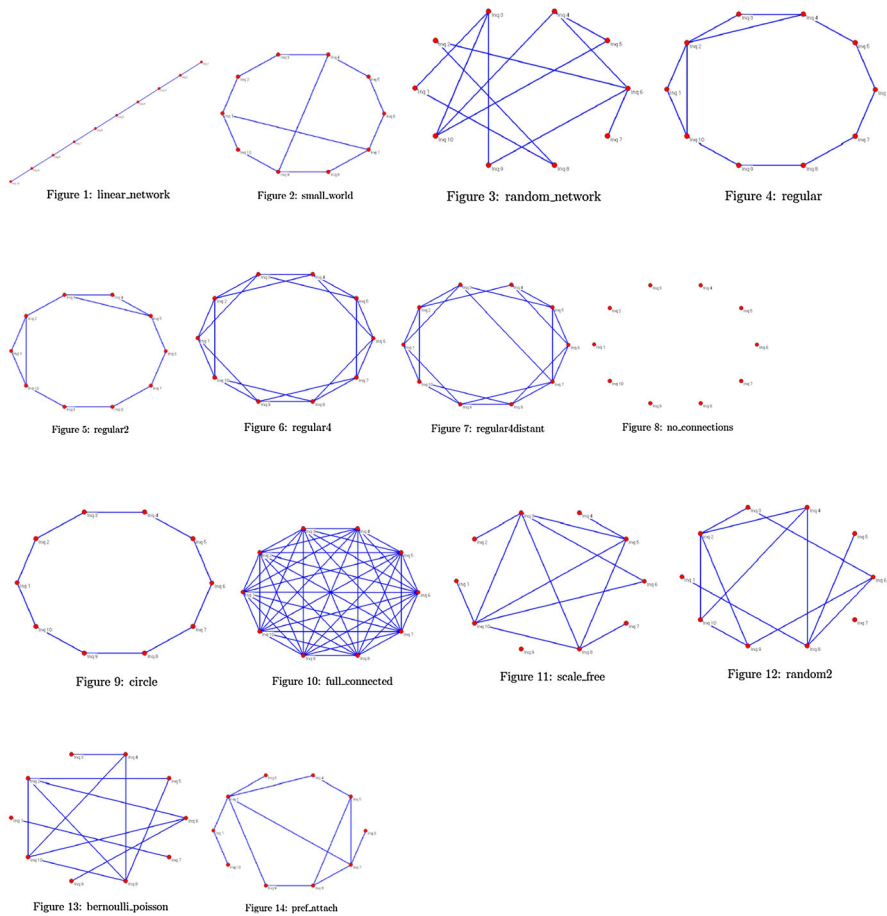


Fig. 2 Networks studied of size 10

structures could be relevant, such as the small world network in which strangers are being linked by a short chain of acquaintances (see Fig. 2).

The fact that parameter values for reliability of inquiry and initial inquiry trust are selected (independently) from the same distribution implies that agents are initially reasonably well calibrated regarding their trust in their respective outside source. Since trust is dynamically updated in the model while the reliability of the outside source remains fixed, the degree of calibration may, and typically does, change in the course of inquiry and deliberation.

Networks were selected for inclusion in the study on the basis of prominence in the literature. Thus, we included all networks in the aforementioned studies by Zollman (2007) and Mason et al. (2007). All in all, 36 networks of size 10, 15 and 18 were included. The networks of size 10 are listed in Fig. 2.

In each case, 10,000 variations of the background parameters (trust, reliability etc.) were studied within the boundaries set by the normality constraints. Each network

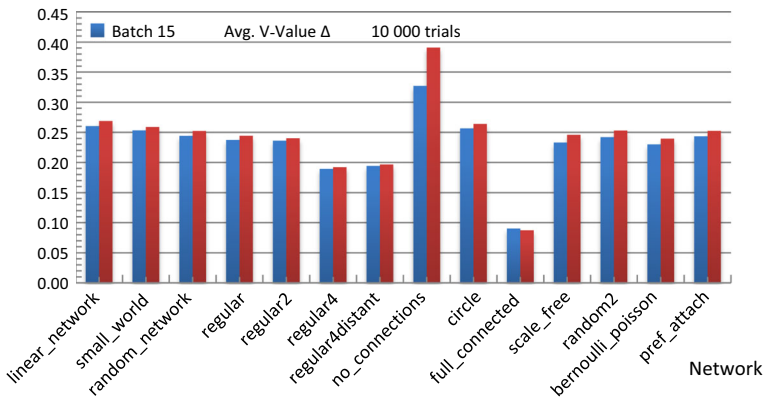


Fig. 3 Increase in V-value as a function of network structure for networks of size 10

deliberation ran for 15 or 30 steps during which inquirers could inquire or communicate. The results to be presented are the average results over these 10,000 runs of the same network structure. The confidence level was 95%, with possible error in the third decimal meaning that visible differences are statistically significant in the figures below.

We have collected further details and background information in several appendices. “Appendix A” contains pictures of all networks included in our study. “Appendix B” contains sample Laputa output in single network mode, and “Appendix C” sketches the derivations of the Laputa updating rules for credence and trust. Finally, “Appendix D” defines and explains the network metrics used.

3 Results

We computed the V-value associated with each network structure against the backdrop of our normality assumptions. The results in terms of increase in V-value for networks of size 10 are displayed in Fig. 3. Blue bars signify results for 15 step simulations and red bars the corresponding results for 30 step simulations.

Combining Figs. 2 and 3, we may conclude that greater connectivity means less V-value. Thus, the fully connected network gives rise to less increase in V-value than, say, the scale free network. On the other hand, more connected networks converge more quickly on a stable state as can be visually confirmed from Fig. 3 by comparing the difference between the blue and corresponding red bar. The smaller the difference is, the quicker the network reaches a stable state. For instance, the regular4distant network converges rapidly whereas the no-connections network continues to improve significantly after 15 steps. Since speed of convergence was not the focus of our study we did not study it systematically. As we mentioned in the introduction, these results are in line with conclusions reached in Zollman (2007) and Mason et al. (2008).

A further observation is that several networks have the same degree of connectivity and yet they differ regarding V-value. This holds for the small world network which is

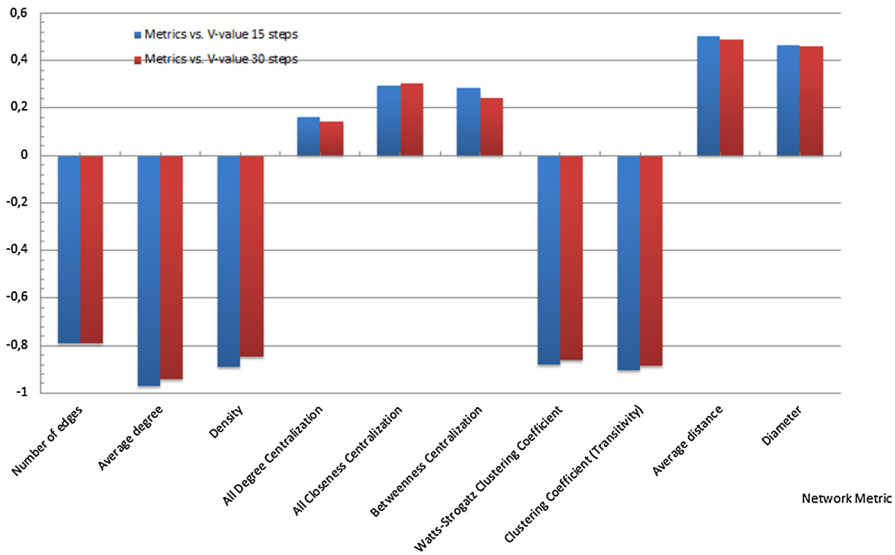


Fig. 4 Correlations between networks metrics and V-values

V-better than the regular network which in turn is V-better than the regular2 network. Similarly, regular4distant performs better regarding V-value than regular4. We may conclude that something other than connectivity is playing a role in determining the V-value of a given network. To find out what is driving these results we studied the correlations between V-value and a collection of prominent network metrics across all 36 networks, drawing on influential work in network analysis (Borgatti 2005; Easley and Kleinberg 2010; Freeman 1977, 1979; Jackson 2010; Milgram 1967; Newman 2010; De Nooy et al. 2011; Watts 1999; Watts and Strogatz 1998). See “Appendix D” for details about these metrics and what they mean. The results are shown in Fig. 4. As before blue bars are results from 15 simulations steps, and red bars are results from 30 simulation steps.

As shown in Fig. 4, we observed positive correlations between V-value and all degree centralization, all closeness centralization, betweenness centralization, average distance and diameter. We registered negative correlations between V-value and number of edges, average degree, density, Watts-Strogatz clustering coefficient and clustering coefficient (transitivity).

Many of the metrics are highly correlated. We thus conducted hierarchical, step-wise regressions in order to identify those metrics that explained unique proportions of the variance in collective competence. On this analysis, only average degree and the clustering coefficient make unique contributions in accounting for differences in veritistic value across all networks studied. More precisely, the following conclusions could be established:

1. Average degree explains 90% of the variation in V-value.
2. A combined model of average degree and clustering coefficient is the best model accounting for 96% of the variation in V-value.

Average degree is the average number of nodes a given node is connected to. This is a measure of connectivity similar to density. The clustering coefficient can be grasped by noting that for a given node we can ask how many of its neighbors are connected themselves. The coefficient now measures the actual number of such “triangles” relative to all possible ones. See “Appendix D” for a definition.

We may conclude that the difference in V-value between networks of the same connectivity (average degree) comes mainly from clustering. The conclusion can quickly be checked by observing that the networks of the same connectivity that we found to be V-better are also less clustered. For example, the small world network is less clustered than the regular network which in turn is less clustered than the regular2 network, and so on. That connectivity and clustering are the driving forces behind our results were confirmed in a further study of larger networks involving seven networks of size 100 and seven networks of size 150 similar to some of the networks included in our main study. Again, we found that networks with a higher average degree promote V-value to a lesser degree and that among networks having the same average degree those that are more clustered perform worse.

4 Discussion

The question remains why we get the results that we get. Why are connectivity and clustering detrimental to collective competence in our study? Note that in our model, agents are assumed to be initially more or less undecided: the initial credence in p was determined by a normal distribution with expected value 0.5 and standard deviation 0.1. Hence there is a fair chance that a majority of inquirers in the network initially tend to believe, falsely, that not- p is the case. The higher the connectivity in the network, the more the misled majority can drag down the whole, or parts of, the network. By Table 1, mechanisms of trust consolidate this phenomenon by strengthening trust within groups of like-minded, and lowering trust in agents delivering belief-contravening (unexpected) information—whether it comes from within or outside the network. A less connected network, by contrast, is better equipped to recover from an unfortunate selection of initial degrees of belief due to the assumed independence and relative reliability of the outside sources.

Bala and Goyal (1998), using a different Bayesian model, observed that more connectivity may have detrimental effects on group competences due to the fact that “more informational links can increase the chances of a society getting locked into a sub-optimal action” (609). Thus, there is reason to think that our proposed explanation may capture a general connectivity effect which is not an artifact of our particular model.

We hypothesize that clustering can be harmful for similar reasons. A cluster which is initially on the wrong track can reinforce itself through internal communication, locking into a false belief. Internal trust turns the cluster into a group of “conspiracy theorists”. This is presumably why the mere rewiring of one of the links in a cluster (as in transition from regular4 to regular4 distant) can have a beneficial effect even though connectivity stays the same.

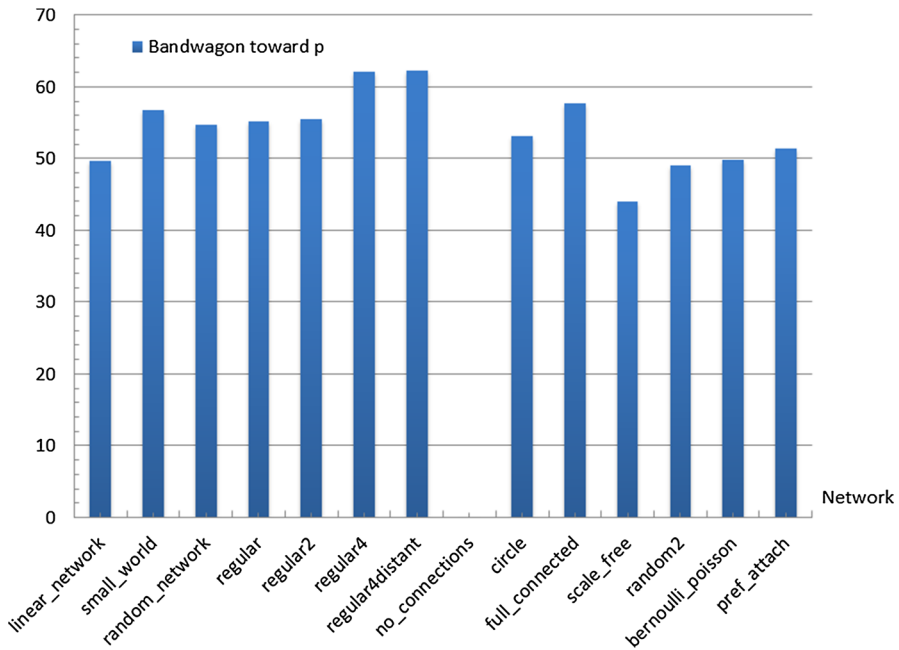


Fig. 5 Amount of good bandwagoning (towards the truth) for different networks

To test these hypotheses we studied the bandwagon effect for various network types, by which we mean the percent of all updates where, as a result of communication from others, an agent's degree of belief has been changed in the opposite direction from her own opinion or information from her outside source. Bandwagoning thus means that you are led to believe something due to social influence that runs counter to your personal information or opinion. As such it is a neutral phenomenon from an epistemological perspective. What matters is whether your peers take you in the *right* direction. Hence, bandwagoning toward p (true) is good, whereas bandwagoning toward $\text{not-}p$ (false) is bad.

Now if our hypotheses are true, then (i) highly connected networks should have some more good, and a lot more bad, bandwagoning, and (ii) more clustered networks should have more bad bandwagoning given same connectivity. Both these predictions turn out to hold in our study, as shown in Figs. 5 and 6.

Thus, a highly connected network like the fully connected network has some more good bandwagoning but a lot more bad bandwagoning than a less connected network such as the circle. Moreover, among networks of the same connectivity, the more clustered ones have more bad bandwagoning. For instance, the regular2 network has more bad bandwagoning than the small world network. In fact, the regular2 network has less good bandwagoning than the small world network as well. At any rate, differences in bad bandwagoning are more pronounced than difference in good bandwagoning for networks of the same connectivity.

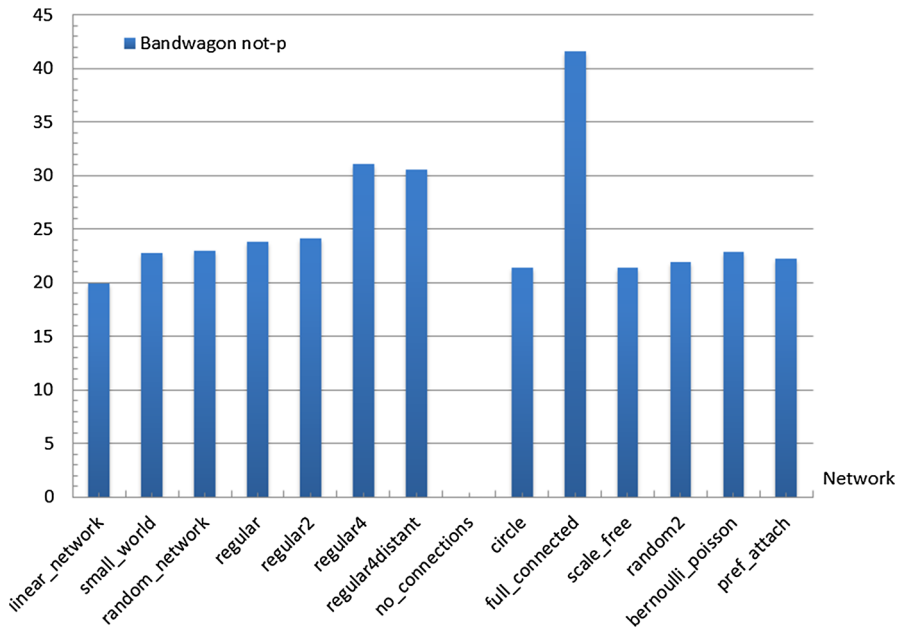


Fig. 6 Amount of bad bandwagoning (towards error) for different networks

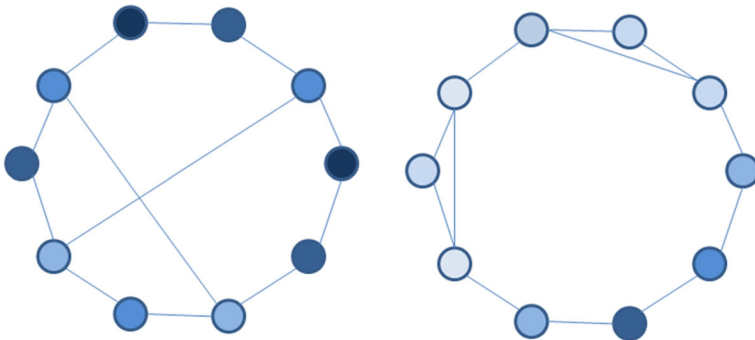


Fig. 7 Final credence in the truth for small world network (left) and regular2 network (right)

Our study of bandwagoning effects in networks supports the truth of our hypotheses, albeit in an indirect, global way. To get a more direct or local sense of what is going on, we zoomed in on the individual nodes in a network to see how clustering affects the agents that occupy the corresponding network positions. We compared two networks from this perspective: small world and regular2. Figure 7 shows the final credences in p for the various network positions. The credences are averages over 10,000 simulations, each simulation running for 15 rounds using the same parameter distributions as before. A fuller color means higher final credence in the true proposition p .

As shown in Fig. 7, agents occupying positions in the clusters in the regular2 network end up with a relatively low credence in the truth, which supports our hypothesis

that clusters have a tendency to reinforce and consolidate false belief. Agents not occupying cluster positions do significantly better. In the less-clustered small world network, differences in outcome between network positions are less salient, although more connected positions are slightly less advantageous than less connected ones. A more detailed study of the effects of network structure on agents occupying individual network position is planned for a future article.

Finally, the fact that the best network is in a sense the “empty” network admittedly renders the rest of our analysis somewhat hollow. Why bother figuring out which among many different networks is better or worse, when keeping people isolated is best? Our first reply is there are many different reasons why people hook up in networks. Improving one’s own epistemic position is surely one of them, but hardly the only one, as the activity on any online social network amply illustrates. Hence, we would expect a network structure in many cases to be given partly by non-epistemic factors, such as a social impulse to communicate. Our model contributes to the tool box that can be used to evaluate an existing network and its variations from a purely epistemic standpoint. Second, the time perspective used in the present study was that of medium to longer term (15 and 30 step simulations, respectively). Preliminary simulations show that connectivity is more attractive and can in fact improve V-performance in shorter simulations. A more extensive investigation into this phenomenon and its causes would require another article. Finally, as confirmed in Angere and Olsson (2017), density becomes V-attractive in Laputa if constraints are introduced that preclude agents from repeating information in the absence of new information from the outside source or other agents. Hence, the simple model used in the present paper corresponds to the case in which agents are free to “spam” the network with repeated messages without having received new evidence in-between—a situation not too unlike that holding in online social networks. A further interesting question, also left for future investigation, is what the correlation between various network metrics, on the one hand, and V-value, on the other, looks like once these “quality constraints” are imposed on communication.

5 Conclusions

We addressed via agent-based computer simulations the extent to which the patterns of connectivity within our social networks affect the likelihood that network peers converge on a true opinion on an issue regarding which they are initially more or less undecided. We explored a wide range of network structures and provided a detailed statistical analysis into the exact contribution of various network metrics to collective competence. Moreover, unlike other similar agent-based models the framework used in this article incorporates a more fine-grained and, we believe, realistic representation of belief and, in particular, trust, where the latter is dynamically updated as agents continuously receive information from their network peers, and the framework also allows for agents to receive information continuously from outside the network.

We found that 96% of the variation in collective competence across different networks can be attributed to differences in amount of connectivity (average degree) and clustering. Both these factors are in our model negatively correlated with collective

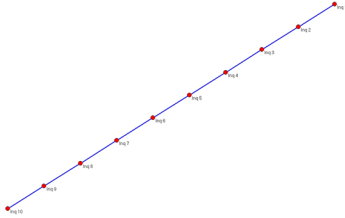
competence. We explained these facts by reference to the increased risk of the group wholly or partly locking into a false belief in a highly connected or clustered network. Our hypotheses were corroborated by observing that connectivity and clustering covary with what we called bad bandwagoning. In other words, initially undecided agents in a tightly connected or clustered network are more likely eventually to have their true personal information or opinion overridden by false group opinion. To be sure, they are more likely to have their false personal information or opinion overridden by true group opinion as well, but this positive effect is less pronounced and also not without exceptions.

By zooming in on individual network positions in two of the studied networks we were able to observe how agents occupying network positions in a cluster ended up with a relatively low average credence in the truth following inquiry and deliberation. Agents not occupying cluster positions did significantly better. In a less clustered network differences in final degrees of belief between network positions are less salient, although our study indicated that more connected network positions are slightly less advantageous than less connected ones. In highlighting and explaining the collective risks which are involved in connectivity and clustering our study suggests that popular belief in the virtues of the network society should give way for a more nuanced picture which takes into account negative effects on the truth tracking properties of networks.

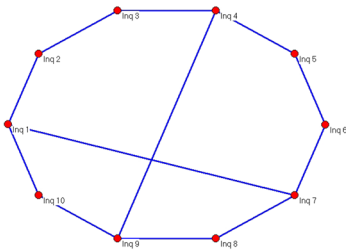
Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Pictures of the networks

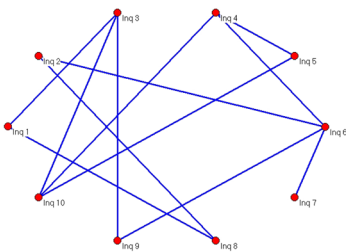
linear_network:



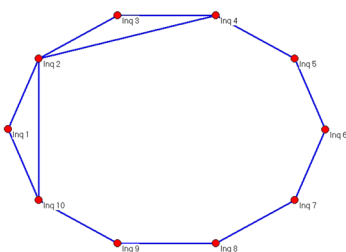
small_world:



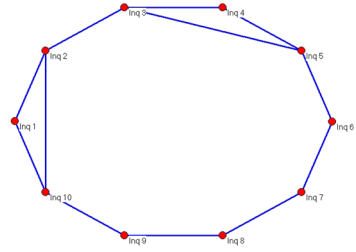
random_network:



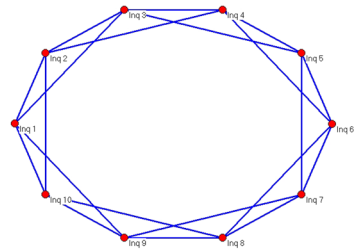
regular:



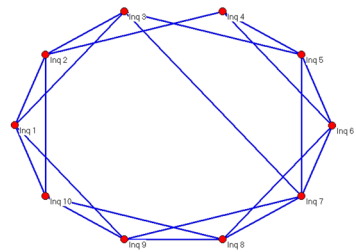
regular:



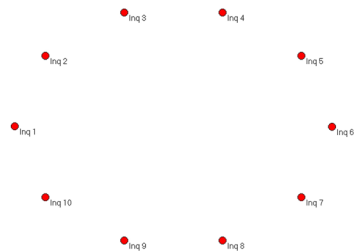
regular4:



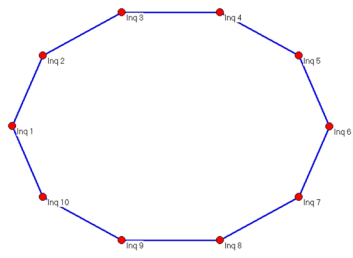
regular4distant:



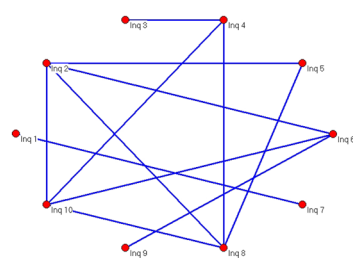
no_connections:



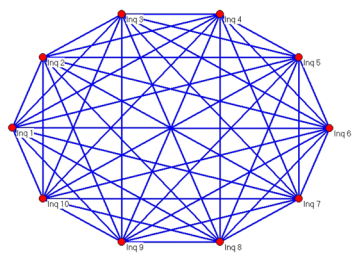
circle:



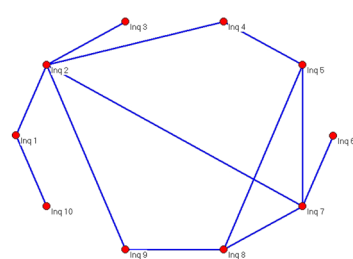
bernoulli_poisson:



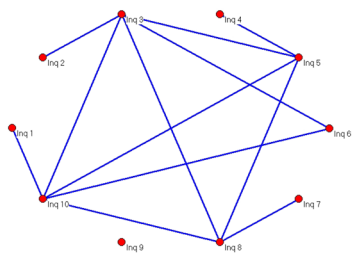
full_connected:



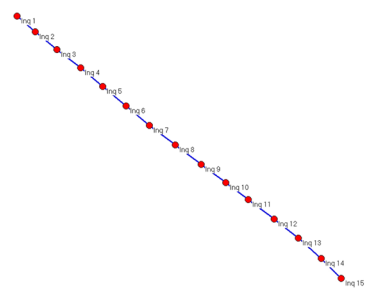
pref_attach:



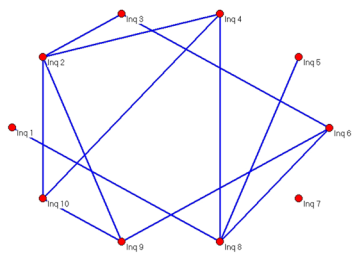
scale_free:



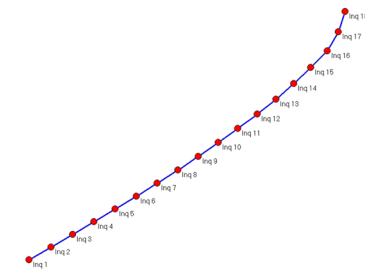
linear_network15:



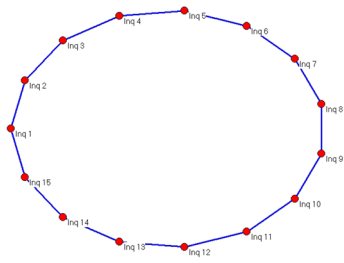
random2:



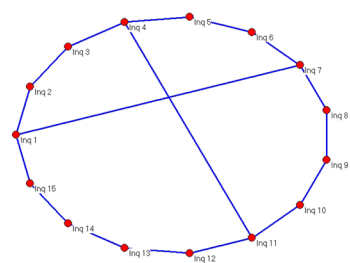
linear_network18:



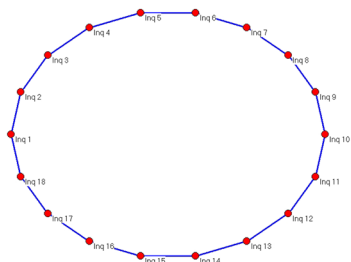
circle15:



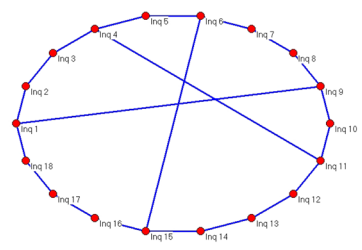
small_world15:



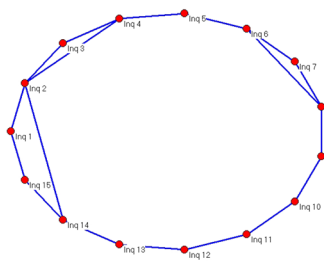
circle18:



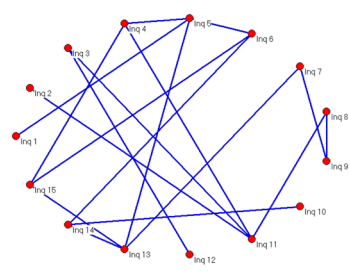
small_world18:



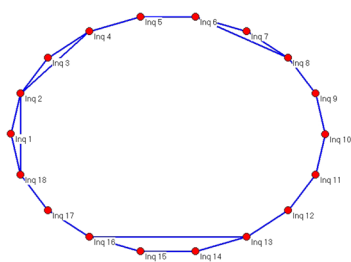
regular15:



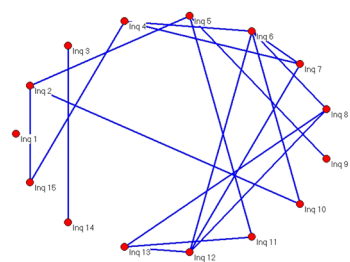
random15:



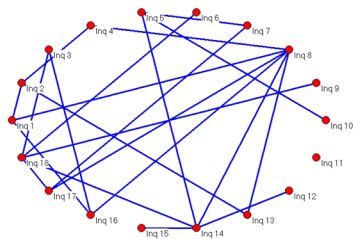
regular18:



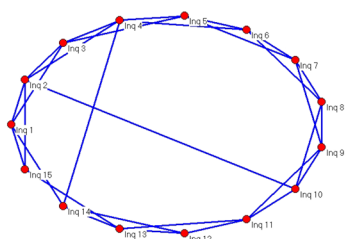
random15_2:



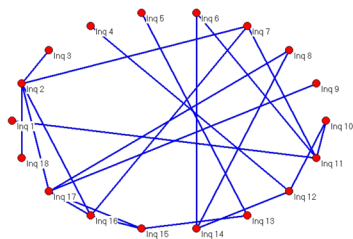
random18:



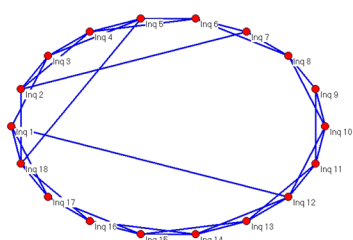
regular4distant_15:



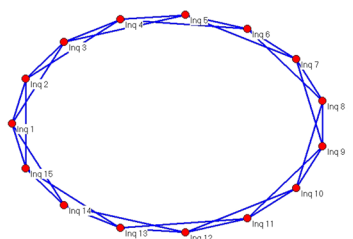
random18_ 2:



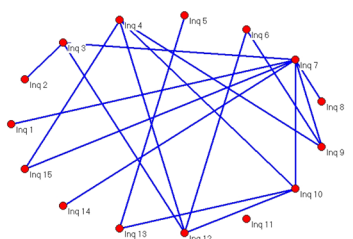
regular4distant_18:



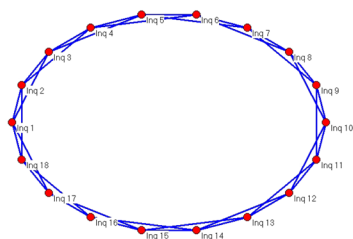
regular4_15:



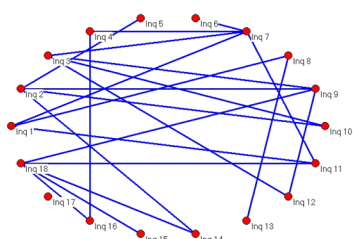
bernoulli_poisson15:



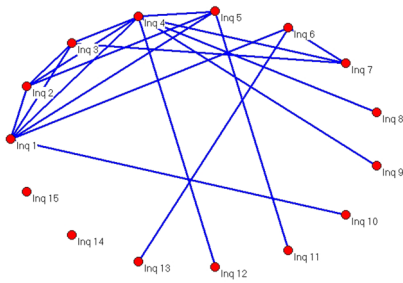
regular4_18:



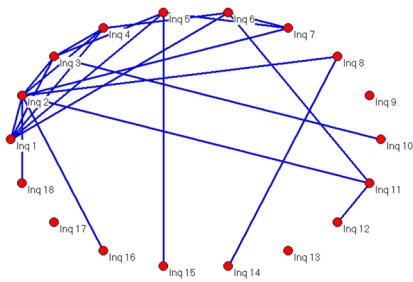
bernoulli_poisson18:



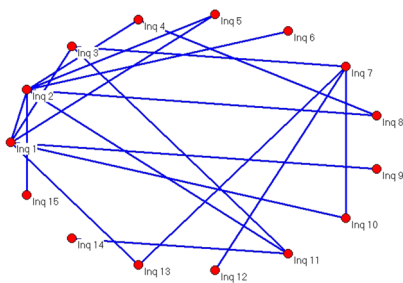
scale_free15:



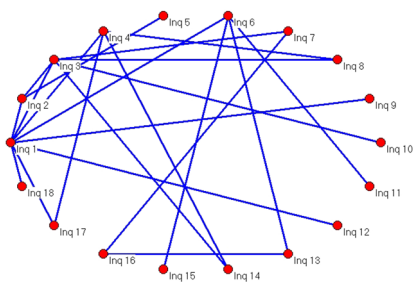
scale_free18:



pref_attach15:

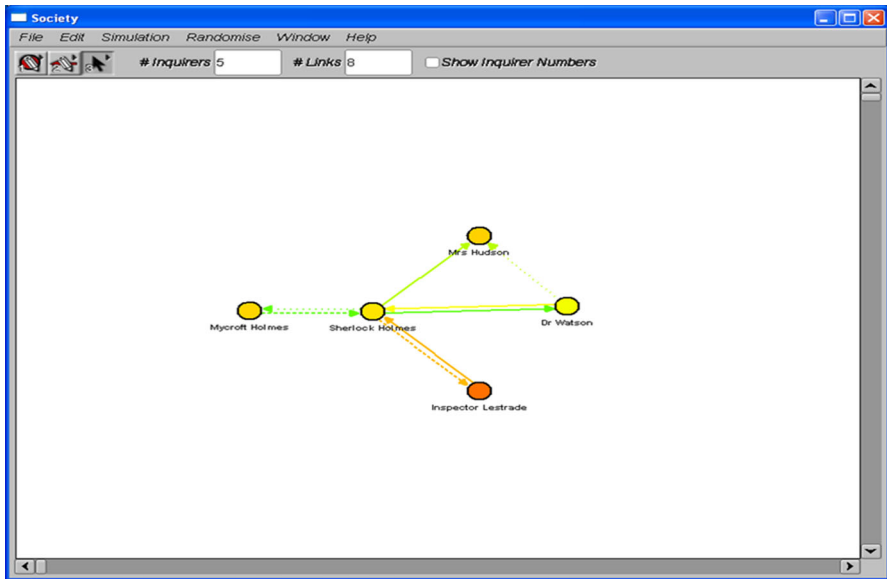


pref_attach18:



Appendix B: Sample Laputa output

We give sample output from Laputa for the “Sherlock Holmes network” depicted below running Laputa in the single network mode. Each time step represents a round in the simulation. What happens during a round is determined by the updating rules in Laputa and the value of the parameters, e.g. inquiry chance and initial degree of belief.



Time: 1

Inquirer Mycroft Holmes heard that p from inquirer Sherlock Holmes, lowering his/her expected trust in the source from 0.513 to 0.513.

This raised his/her degree of belief in p from 0.50000 to 0.51299.

Time: 2

(* Nothing happened *)

Time: 3

Inquirer Mrs Hudson received the result that not- p from inquiry, raising his/her expected trust in it from 0.642 to 0.672.

This lowered his/her degree of belief in p from 0.27923 to 0.17761

Time: 4

Inquirer Sherlock Holmes received the result that p from inquiry, raising his/her expected trust in it from 0.371 to 0.452.

Inquirer Sherlock Holmes heard that not- p from inquirer Prof Moriarty, lowering his/her expected trust in the source from 0.656 to 0.573.

This lowered his/her degree of belief in p from 0.91000 to 0.75815

Time: 5

Inquirer Sherlock Holmes received the result that not- p from inquiry, lowering his/her expected trust in it from 0.452 to 0.414.

This raised his/her degree of belief in p from 0.75815 to 0.79160.

Time: 6

Inquirer Sherlock Holmes heard that not- p from inquirer Prof Moriarty, lowering his/her expected trust in the source from 0.573 to 0.525.

This lowered his/her degree of belief in p from 0.79160 to 0.73859

Inquirer Dr Watson heard that not- p from inquirer Mrs Hudson, lowering his/her expected trust in the source from 0.581 to 0.576.

This lowered his/her degree of belief in p from 0.53000 to 0.44843

...

Appendix C: The mathematics behind Laputa

In this appendix we sketch the derivations of the credence and trust update function in Laputa. These derivations make use of a number of idealizations and technical assumptions. The reader may want to consult Olsson (2013) or Vallinder and Olsson (2014) for more details on the intuitive meaning and justification of these assumptions.

The following assumptions are used in the derivation of the credence update function:

Principal Principle (PP):

$$\begin{aligned} a < C_{\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) < b \\ a < C_{\alpha}^t(S_{\sigma\alpha}^t(\neg p)|S_{\sigma\alpha}^t, \neg p, a < r_{\sigma\alpha} < b) < b \end{aligned}$$

Communication Independence (CI):

$$\begin{aligned} C_{\alpha}^t(S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) &= C_{\alpha}^t(S_{\sigma\alpha}^t(p)) \times C_{\alpha}^t(p) \times C_{\alpha}^t(a < r_{\sigma\alpha} < b) \\ C_{\alpha}^t(S_{\sigma\alpha}^t, \neg p, a < r_{\sigma\alpha} < b) &= C_{\alpha}^t(S_{\sigma\alpha}^t(p)) \times C_{\alpha}^t(\neg p) \times C_{\alpha}^t(a < r_{\sigma\alpha} < b) \end{aligned}$$

Source Independence (SI):

$$\begin{aligned} C_{\alpha}^t\left(\bigwedge_{\sigma \in \Sigma_{\alpha}^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p\right) &= \prod_{\sigma \in \Sigma_{\alpha}^t} C_{\alpha}^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p) \\ C_{\alpha}^t\left(\bigwedge_{\sigma \in \Sigma_{\alpha}^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p\right) &= \prod_{\sigma \in \Sigma_{\alpha}^t} C_{\alpha}^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p) \end{aligned}$$

where $r_{\sigma\alpha}$ is the reliability of the source σ vis-à-vis agent α , $0 \leq a < b \leq 1$, $C_{\alpha}^t(p)$ the credence that agent α assigns to p at time t , $S_{\sigma\alpha}^t(p)$ the proposition that source σ communicates p to agent α at time t , $m_{\sigma\alpha}^t$ is the content of the source σ 's message, and Σ_{α}^t is the set of sources that give information to α at t .

Since the trust function, which plays a crucial part in the model, is continuous, the derivation will sometimes need to take a detour through conditional probability densities rather than the conditional probabilities themselves. We will briefly sketch how this can be done here.

We have so far not been specific about the σ -algebra Z that C_α^t is defined on. Assume that it is product of several such algebras, the first of which is discrete and generated by atomic events such as $p, \neg p, S_{\beta\alpha}(p)$ etc., and the others, which are continuous, are generated by events of the form $a \leq r_{\sigma\alpha} \leq b$. Call the first algebra X and the others $Y_{\sigma_0}, \dots, Y_{\sigma_n}$. It is clear that, as long as time and the number of inquirers are both finite, X will have only finitely many elements. On the other hand, $Y_{\sigma_0}, \dots, Y_{\sigma_n}$ are certainly infinite. As mentioned, we assume that $Z = X \times Y_{\sigma_0} \times \dots \times Y_{\sigma_n}$. Given any source σ_k and time t , we can therefore interpret the part of C_α^t defined on the subalgebra $X \times Y_{\sigma_k}$ of Z as arising from a joint density function $\kappa_{\sigma\alpha}^t(\varphi; x)$ defined through the equation

$$C_\alpha^t(\varphi, a < r_{\sigma\alpha} < b) = \int_a^b \kappa_{\sigma\alpha}^t(\varphi; x) dx$$

Since we have used the comma to represent conjunction earlier in the paper we use a semicolon here to separate the two variables: the first propositional, and the second real-valued. Like τ , this distribution's existence and essential uniqueness are guaranteed by the Radon-Nikodym theorem, and in fact $\tau_{\sigma\alpha}^t$ is the marginal distribution of $\kappa_{\sigma\alpha}^t$ with respect to the reliability variable $r_{\sigma\alpha}$ in question. Since the conditional distribution of a random variable is the joint distribution divided by the marginal distribution of that variable, this means that we have that

$$\kappa_{\sigma\alpha}^t(\varphi|x) = \frac{\kappa_{\sigma\alpha}^t(\varphi; x)}{\tau_{\sigma\alpha}^t(x)}$$

which is what will be used to make sense of what it means to conditionalize on $r_{\sigma\alpha}$ having a certain value rather than merely being inside an interval. Setting $r_{\sigma\alpha} = x$, $a = x - \epsilon$ and $b = x + \epsilon$ in *PP* and *CI* and letting $\epsilon \rightarrow 0$, we get the versions

$$\text{Principal Principle}_0(PP_{lim}) : \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t(p); x) = \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(\neg p)|S_{\sigma\alpha}^t(\neg p); x) = x$$

$$\begin{aligned} \text{Communication Independence}_0(CI_{lim}) : \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p); x) &= C_\alpha^t(S_{\sigma\alpha}^t(p)) \times C_\alpha^t(p) \times \tau_{\sigma\alpha}^t(x) \\ \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(\neg p); x) &= C_\alpha^t(S_{\sigma\alpha}^t(p)) \times C_\alpha^t(\neg p) \times \tau_{\sigma\alpha}^t(x) \end{aligned}$$

We can now proceed with the actual derivation. By conditionalization, we must have that $C_\alpha^t(p)$ is equal to $C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p\right)$. Applying Bayes' theorem and then *SI* to this expression gives

$$\begin{aligned} C_\alpha^t\left(p|\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)\right) &= \frac{C_\alpha^t(p) \times C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p\right)}{C_\alpha^t(p) \times C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p\right) + C_\alpha^t(\neg p) \times C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p\right)} \\ &= \frac{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p)}{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p) + C_\alpha^t(\neg p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p)} \end{aligned}$$

which gives us the posterior credence in terms of the values $C_\alpha^t(S_{\sigma\alpha}^t(p)|p)$ and $C_\alpha^t(S_{\sigma\alpha}^t(\neg p)|\neg p)$. Our next task is thus to derive these expressions. Since $S_{\sigma\alpha}^t(p)$ is equivalent to $S_{\sigma\alpha}^t(p) \wedge S_{\sigma\alpha}^t$, it follows that $C_\alpha^t(S_{\sigma\alpha}^t(p)|p) = C_\alpha^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t|p)$.

Applying first the definition of conditional probability and then the continuous law of total probability, the definition of conditional probability again, and finally CI_{lim} , we get, after some calculations,

$$\begin{aligned} C_{\alpha}^t(S_{\sigma\alpha}^t(p)|p) &= \frac{1}{C_{\alpha}^t(p)} \times C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p) \\ &= C_{\alpha}^t(S_{\sigma\alpha}^t) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p; x) \times \tau_{\sigma\alpha}^t(x) dx \end{aligned}$$

But PP_{lim} ensures that $\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p; x) = x$, so we get

$$C_{\alpha}^t(S_{\sigma\alpha}^t(p)|p) = C_{\alpha}^t(S_{\sigma\alpha}^t) \times \int_0^1 x \times \tau_{\sigma\alpha}^t(x) dx = C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]$$

Parallel derivations give that

$$C_{\alpha}^t(S_{\sigma\alpha}^t(\neg p)|\neg p) = C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}]$$

Now let $\Sigma_{\alpha}^t(p) \subseteq \Sigma_{\alpha}^t$ be the set of sources that give α the message p at t , and let $\Sigma_{\alpha}^t(\neg p) = \Sigma_{\alpha}^t \setminus \Sigma_{\alpha}^t(p)$. Plugging the above expressions into our earlier result gives the sought for expression

$$C_{\alpha}^{t+1}(p) = \frac{C_{\alpha}^t(p) \times \prod_{\sigma \in \Sigma_{\alpha}^t(p)} C_{\alpha}^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p)}{C_{\alpha}^t(p) \times \prod_{\sigma \in \Sigma_{\alpha}^t(p)} C_{\alpha}^t(S_{\sigma\alpha}^t) + C_{\alpha}^t(\neg p) \times \prod_{\sigma \in \Sigma_{\alpha}^t(\neg p)} C_{\alpha}^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p)} = \frac{\gamma}{\gamma + \delta}$$

where

$$\begin{aligned} \gamma &= C_{\alpha}^t(p) \times \prod_{\sigma \in \Sigma_{\alpha}^t(p)} (C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]) \times \prod_{\sigma \in \Sigma_{\alpha}^t(\neg p)} (C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}]) \\ \delta &= C_{\alpha}^t(\neg p) \times \prod_{\sigma \in \Sigma_{\alpha}^t(p)} (C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}]) \times \prod_{\sigma \in \Sigma_{\alpha}^t(\neg p)} (C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]) \end{aligned}$$

For the derivation of the trust update expression we assume PP and CI , but not SI . The function we wish to derive is

$$\tau_{\sigma\alpha}^{t+1}(x) = \kappa_{\sigma\alpha}^t(x|S_{\sigma\alpha}^t(m_{\sigma\alpha}^t))$$

for a source σ of α , and a message $m_{\sigma\alpha}^t$ from that source. Assume that $m_{\sigma\alpha}^t \equiv p$ (the case $m_{\sigma\alpha}^t \equiv \neg p$ is completely symmetrical). Applying the definition of conditional probability, the equivalence $S_{\sigma\alpha}^t \wedge S_{\sigma\alpha}^t(p) \equiv S_{\sigma\alpha}^t(p)$, and the (discrete) law of total probability, we get

$$\begin{aligned} \kappa_{\sigma\alpha}^t(x|S_{\sigma\alpha}^t(p)) &= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p); x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} = \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \\ &= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p; x) + \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, \neg p; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \end{aligned}$$

$$= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t(p), p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), p; x) + \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t(p), \neg p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), \neg p; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))}$$

Now apply PP_{\lim} and CI_{\lim} to the factors in both terms of the numerator, and then again the equivalence $S_{\sigma\alpha}^t \wedge S_{\sigma\alpha}^t(p) \equiv S_{\sigma\alpha}^t(p)$:

$$\begin{aligned}\kappa_{\sigma\alpha}^t(x|S_{\sigma\alpha}^t(p)) &= \tau_{\sigma\alpha}^t(x) \times C_{\alpha}^t(S_{\sigma\alpha}^t) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \\ &= \tau_{\sigma\alpha}^t(x) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t(p))}\end{aligned}$$

We can calculate the denominator in this expression by using the definition of conditional probability and expanding twice using the law of total probability (once using the discrete version, and once using the continuous one):

$$\begin{aligned}C_{\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t) &= \frac{C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t)}{C_{\alpha}^t(S_{\sigma\alpha}^t)} = \frac{C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p) + C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, \neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \\ &= \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, p; x) dx \\ &\quad + \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, \neg p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, \neg p; x) dx\end{aligned}$$

Let us refer to the last expression as ψ . Applying CI_{\lim} , then cancelling, and applying PP_{\lim} , we get

$$\begin{aligned}\psi &= C_{\alpha}^t(p) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(x) dx \\ &\quad + C_{\alpha}^t(p) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, \neg p; x) \times \kappa_{\sigma\alpha}^t(x) dx \\ &= C_{\alpha}^t(p) \times \int_0^1 x \times \tau_{\sigma\alpha}^t(x) dx + C_{\alpha}^t(\neg p) \times \int_0^1 \bar{x} \times \tau_{\sigma\alpha}^t(x) dx \\ &= C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]\end{aligned}$$

Putting it all together, we finally arrive at the updating rule for trust:

$$\tau_{\sigma\alpha}^{t+1}(x) = \tau_{\sigma\alpha}^t(x) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]}$$

Appendix D: Network metrics used

In this appendix we will properly define the network measures we have calculated for the various networks in our study. These measures are all common measures from

the network literature [see Easley and Kleinberg (2010), Jackson (2010) or Newman (2010)]. The terminology used in this section will mostly be borrowed from Jackson (2010), but as the measures for the particular networks were calculated using the program Pajek (De Nooy et al. 2011), we will deviate from Jackson (2010) whenever the measures are calculated in a different way in Pajek.

By a network we will understand an undirected graph (N, g) , where $N = \{1, 2, \dots, n\}$ is the set of nodes also referred to as vertices, individuals, or inquirers. g is a set of pairs (i, j) specifying which links between nodes are present in the network. We will also write ij for (i, j) and if $ij \in g$ then we will say that there is a link/edge/tie between i and j . As networks are undirected graphs we will not distinguish between ij and ji (i.e. $ij \in g$ will be equivalent to $ji \in g$). In the following (N, g) will refer to a given arbitrary network with $N = \{1, 2, \dots, n\}$. The *neighborhood* $N_i(g)$ of a node i is the set of nodes that i is linked to, that is $N_i(g) = \{j | ij \in g\}$. The degree of a node i , denoted by $d_i(g)$, is the total number of nodes i is linked, in other words, $d_i(g) = \#N_i(g)$, where $\#A$ denotes the cardinality of the set A .

Our first measure is the *average degree*, which says something about how connected each inquirer is on average, and is defined by:

$$d^{Avg}(g) = \frac{1}{n} \sum_{i=1}^n d_i(g).$$

A related measure is the *density* of a network. The density of a network is the fraction of links actually present in the network relative to the number of possible links. For a network with n nodes, the number of all possible links is $n(n-1)/2$ (as the network is undirected). Thus the density is given by²:

$$den(g) = \frac{\#g}{n(n-1)/2}.$$

Note that

$$den(g) = \frac{d^{Avg}(g)}{n-1},$$

and thus, average degree and density are highly correlated—even perfectly correlated if the size of the networks is kept fixed. As exhaustively discussed in the introduction, the significance of a network's density on its potential as an information passing structure, has been widely discussed in the literature.

While average degree and density say something about how connected a network is, there are much more to a network's structure than captured by these measures. Another structural property one could consider is the distances between nodes (how many links one has to pass through to reach one node from another), which may say something about how fast things such as diseases or information can spread through a network. It turns out that the distances between nodes are surprisingly small in real networks, which is known as the "Small-world phenomenon" (or the "six degrees of

² Here we only count either $ij \in g$ or $ji \in g$, but not both!

separation” or the “Kevin Bacon effect” (Watts 1999)) and goes back to a famous experiments by Milgram (1967). Let us turn to the formal definitions.

A *path* from a node i to another node j is a sequence of distinct nodes i_1, i_2, \dots, i_m such that $i_1 = i$, $i_m = j$, and $i_k i_{k+1} \in g$ for all $k \in \{1, 2, \dots, m-1\}$. In other words, there is a path from i to j if one can reach j from i by following a sequence of distinct links in the network. The length of such a path is the number of links in it (i.e. $m-1$). A *shortest path* from i to j is a path from i to j such that there are no other path from i to j with a shorter length. The *distance* between two nodes i and j is the length of a shortest path between them (if there is any path between i and j at all) and will be denoted by $dist(i, j)$. We say that two nodes are connected if there is a path (and thereby necessarily a shortest path) between them. A network is *connected* if every pair of distinct nodes i and j are connected. Traditionally, the *average distance* of a network is the average distance of any two nodes in it, that is

$$\frac{1}{n(n-1)} \sum_{i,j \in N} dist(i, j).$$

However, since there might not always be a path between two nodes i and j (if the network is not connected), the average distance is calculated a little bit different in Pajek. Let,

$$A := \{(i, j) | i, j \in N, i \neq j, \text{ and } dist(i, j) \text{ is defined}\}.$$

Then, the average distance, $dist^{Avg}(g)$, is calculated in Pajek in the following way:

$$dist^{Avg}(g) = \frac{1}{\#A} \sum_{(i,j) \in A} dist(i, j).$$

The diameter of a network is the length of the longest shortest part in it:

$$diameter(g) = \max_{i,j \in N} dist(i, j).$$

If the network is not connected, Pajek calculates the above formula as if $dist(i, j)$ is 0 for nodes i and j that are not connected. Thereby, the diameter is always the maximum of all shortest paths.³ Note that the diameter of a network puts an upper bound on the average distance of the network, but the diameter can sometimes be significantly larger than the average distance.

A node’s particular position in a network can be important and centralization measures are all trying to measure this effect. We will consider three different measures of centrality that are widely used in social network analysis. As such these measures are local to nodes of a network, but Pajek also provides global “summations” of the local

³ Jackson (2010) defines $dist(i, j)$ to be infinite if i and j are not connected. Thus, disconnected networks have infinite diameters according to Jackson (2010). Alternatively, one could report as the diameter of a disconnected network, the diameter of the largest connected component of it (i.e. the largest subset of nodes that forms a connected network).

measures that we will use. The first simple measure is *degree centralization*, which measures the degree of a node relative to the size of the network, i.e. for a node i the degree centrality of i is defined by:

$$Ce_i^D(g) = \frac{d_i(g)}{n-1}.$$

Pajek provides a measure for the overall degree centrality of a network (“All Degree Centralization”), $Ce^D(g)$, by the following calculation:

$$Ce^D(g) = \frac{\sum_{i \in N} (Ce_*^D(g) - Ce_i^D(g))}{n-2},$$

where $Ce_*^D(g)$ is the maximum of the individual degree centralities $Ce_i^D(g)$.

Degree centrality captures some form of importance of centrality of nodes, however, it is often way to simple. For instance, a node with only two links might connect two otherwise separated part of a network and thereby play a central role if information (or other things) has to pass from the one part to the other. *Betweenness centrality* is a measure, initially defined by Freeman (1977), which is an attempt to capture such potential control over communication. Specifically, betweenness centrality measures how many shortest paths (between other nodes) a given node lies on. Formally, for a node i the betweenness centrality of i is defined by:

$$Ce_i^B(g) = \sum_{k, j \in N, k \neq j, k \neq i, j \neq i} \frac{P_i(kj)/P(kj)}{(n-1)(n-2)/2},$$

where $P_i(kj)$ is the number of shortest paths between k and j that include i and $P(kj)$ is the total number of shortest paths between k and j . Again, Pajek provides a measure for the overall betweenness centrality of a network (“Betweenness Centralization”), $Ce^B(g)$, in the following sense:

$$Ce^B(g) = \frac{\sum_{i \in N} (Ce_*^B(g) - Ce_i^B(g))}{n-1},$$

where $Ce_*^B(g)$ is the maximum of the individual betweenness centralities.

While betweenness centrality may say something about who controls the flow of information in a network, *closeness centrality* is another measure that may say something about how fast or how far information from a node spreads. Formally, closeness centrality measures how close a node is to all the other nodes of the network. For a node i , the closeness centrality of i is defined as:

$$Ce_i^C(g) = \frac{n-1}{\sum_{j \in N, j \neq i} dist(i, j)}.$$

If i is an isolated node in the network, i.e. $d_i(g) = 0$, the convention is that $Ce_i^C(g)$ is 0. If the network is connected, Pajek provides a measure for the overall closeness

centrality of a network (“All Closeness Centralization”), $Ce^C(g)$, by the following calculation:

$$Ce^C(g) = \frac{\sum_{i \in N} (Ce_*^C(g) - Ce_i^C(g))}{(n-1)(n-2)/(2n-3)}.$$

where $Ce_*^C(g)$ is the maximum of the individual closeness centralities. For more on centrality measures and their use see Freeman (1979) and Borgatti (2005) or any of the textbooks referenced in the beginning of this section.

A final class of measures that we will consider is clustering or transitivity measures. Intuitively such measures say something about how likely it are that any two of my friends are friends themselves. It turns out that in many real social networks clustering is much higher than in most random networks (Watts and Strogatz 1998). Formally, for a given node we can ask how many of its neighbors are connected themselves, which give rise to the following *individual clustering* measure:

$$Cl_i(g) = \frac{\#\{jk \in g | k \neq j \text{ and } j, k \in N_i(g)\}}{\#\{jk | k \neq j \text{ and } j, k \in N_i(g)\}} = \frac{\#\{jk \in g | k \neq j \text{ and } j, k \in N_i(g)\}}{d_i(g)(d_i(g)-1)/2}.$$

Taking the average of this measure results in a measure of *average clustering*⁴ of an entire network:

$$Cl^{Avg}(g) = \sum_{i=1}^n \frac{Cl_i(g)}{n}.$$

This measure of average clustering is referred to in Pajek as “Watts-Strogatz Clustering Coefficient” as it was first proposed by Watts and Strogatz (1998). Note that, commonly, $Cl_i(g)$ is taken to be 0 if the neighborhood of i only contains one or zero nodes (see Jackson (2010) or Newman (2010)). Conversely, Pajek takes $Cl_i(g)$ to be plus infinity in this case and does not include the node when calculating the average. In other words, in Pajek:

$$Cl^{Avg}(g) = \sum_{i \in N, d_i(g) > 1} \frac{Cl_i(g)}{\#\{j | d_j(g) \geq 1\}}.$$

An alternative way of measuring clustering is by reporting the actual number of “triangles” relative to all possible triangles:

$$Cl(g) = \frac{\sum_{i=1}^n \#\{jk \in g | k \neq j \text{ and } j, k \in N_i(g)\}}{\sum_{i=1}^n \#\{jk | k \neq j \text{ and } j, k \in N_i(g)\}}.$$

This measure is usually referred to as *overall clustering*, while in Pajek, the measure is referred to as “Clustering Coefficient (Transitivity)”. Note that, overall clustering and average clustering can be different for a particular network.

⁴ According to Newman (2010), the measure for individual clustering was first introduced by Watts (1999), while the idea of measuring average clustering was proposed by Watts and Strogatz (1998).

References

- Angere, S., & Olsson, E. J. (2017). Publish late, publish rarely!: Network density and group performance in scientific communication. In T. Boyer, C. Mayo-Wilson, & M. Weisberg (Eds.), *Scientific collaboration and collective knowledge*. Oxford: Oxford University Press.
- Bala, V., & Goyal, S. (1998). Learning from neighbours. *Review of Economic Studies Limited*, 65, 595–621.
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55–71.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78, 1–3.
- Collins, P. J., Hahn, U., von Gerber, Y. & Olsson, E. J. (2018). The bi-directional relationship between source characteristics and message content. *Frontiers in Psychology*, 9, 18.
- De Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with pajek*. Cambridge: Cambridge University Press.
- DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345), 118–121.
- Douven, I., & Kelp, C. (2011). Truth approximation, social epistemology, and opinion dynamics. *Erkenntnis*. <http://link.springer.com/article/10.1007/s10670-011-9295-x/fulltext.html>.
- Easley, E., & Kleinberg, J. (2010). *Networks, crowds, and markets: reasoning about a highly connected world*. Cambridge: Cambridge University Press.
- Fallis, D. (2007). Attitudes toward epistemic risk and the value of experiments. *Studia Logica*, 86, 215–246.
- Frey, D., & Šešelja, D. (2018). Robustness and idealizations in agent-based models of scientific interaction. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axy039>.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- Freeman, L. C. (1979). Centrality in social networks: conceptual clarification. *Social Networks*, 1, 215–239.
- Goldman, A. I. (1999). *Knowledge in a social world*. Oxford: Clarendon Press.
- Golub, Benjamin, & Jackson, M. O. (2010). Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1), 112–149.
- Hahn, U., Hansen, J. U., & Olsson, E. J. (in preparation). Information networks, truth, and value.
- Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence: Models, analysis, and simulations. *Journal of Artificial Societies and Social Simulation* 5. <http://jasss.soc.surrey.ac.uk/5/3/2.html>.
- Hegselmann, R., & Krause, U. (2006). Truth and cognitive division of labor: first steps towards a computer aided social epistemology. *Journal of Artificial Societies and Social Simulation* 9. <http://jasss.soc.surrey.ac.uk/9/3/10.html>.
- Isenberg, D. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50(6), 1141–1151.
- Jackson, M. O. (2010). *Social and economic networks*. Princeton: Princeton University Press.
- Jönsson, M., Hahn, U., & Olsson, E. J. (2015). The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition*, 142, 191–204.
- Joyce, J. M. (1998). A Nonpragmatic vindication of probabilism. *Philosophy of Science*, 65, 575–603.
- Kopec, M. (2012). We ought to agree: A consequence of repairing goldman's group scoring rule. *Episteme*, 9(2), 101–114.
- Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Computer and Information Science Faculty Publications*, paper 1. <http://hdl.handle.net/2047/d20000313>.
- Maher, P. (1993). *Betting on theories*. Cambridge: Cambridge University Press.
- Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating social influence processes: Dynamic, multi-directional flows of influence within social networks. *Personality and Social Psychology Review*, 11, 279–300.
- Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3), 422–433.
- Masterton, G., & Olsson, E. J. (2014). Argumentation and belief updating in social networks: A Bayesian model. In E. Fermé, D. Gabbay, & G. Simari (Eds.), *Trends in belief revision and argumentation dynamics*. Cambridge: College Publications.
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 2, 60–67.
- Newman, M. (2010). *Networks: An introduction*. Oxford: Oxford University Press.
- Olsson, E. J. (2011). A simulation approach to veritistic social epistemology. *Episteme*, 8(2), 127–143.

- Olsson, E. J. (2013). A bayesian simulation model of group deliberation and polarization. In F. Zenker (Ed.), *Bayesian argumentation, Synthese Library* (pp. 113–134). New York: Springer.
- Olsson, E. J., & Vallinder, A. (2013). Norms of assertion and communication in social networks. *Synthese*, 190, 1437–1454.
- Schurz, G. (2009). Meta-induction and social epistemology: Computer simulations of prediction games. *Episteme*, 6(2), 200–220.
- Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy*, 10(2), 175–195.
- Thorn, P., & Schurz, G. (2012). Meta-induction and the wisdom of crowds. *Analyse & Kritik*, 34(2), 339–366.
- Vallinder, A., & Olsson, E. J. (2014). Trust and the value of overconfidence: A bayesian perspective on social network communication. *Synthese*, 191, 1991–2007.
- Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness*. Princeton: Princeton University Press. ISBN 9780691117041.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393, 440–442.
- Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74(5), 574–587.

Affiliations

Ulrike Hahn¹ · Jens Ulrik Hansen² · Erik J. Olsson² 

¹ Department of Psychological Science, Birkbeck University of London, London, UK

² Department of Philosophy, Lund University, Lund, Sweden